(19)日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号 特開2001-243256 (P2001-243256A)

(43)公開日 平成13年9月7日(2001.9.7)

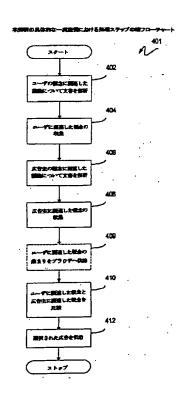
(51) Int.Cl. ⁷	識別記号	F I デーマコート*(参考)
G06F 17/30	3 4 0	G 0 6 F 17/30 3 4 0 A 5 B 0 4 9
	170	170Z 5B075
	2 1 0	2 1 0 A
13/00	5 4 0	13/00 5 4 0 P
17/60	3 2 6	17/60 3 2 6
		審査請求 未請求 請求項の数16 OL (全 36 頁
(21)出願番号	特願2000-347730(P2000-347730)	(71)出願人 000006747
		株式会社リコー
(22)出顧日	平成12年11月15日(2000.11.15)	東京都大田区中馬込1丁目3番6号
		(72)発明者 ジャメイ グラハム
(31)優先権主張番号	483092	アメリカ合衆国,カリフォルニア 94025,
(32)優先日	平成12年1月14日(2000.1.14)	メンロ・パーク, サンド・ヒル・ロード
(33)優先権主張国	米国 (US)	2882番, スイート115, リコーコーポレ
	•	ーション内
		(74)代理人 100070150
		弁理士 伊東 忠彦 (外1名)
		最終頁に統

(54) 【発明の名称】 ウェブ広告に基づく内容表示方法及び装置と、内容表示プログラム

(57)【要約】

【課題】 本発明は、インターネット上で閲覧者の関心 に基づいてオンライン広告を閲覧者に配布する的を絞っ た販売システムの提供を目的とする。

【解決手段】 本発明の一実施例は、ユーザの関心の概念と、広告主の関心の概念と、閲覧中の文書のマッチングを用いて、文書の閲覧者へ広告を的確に送る。ブラウザで閲覧されているページ毎に前後関係に対応した広告を生成し、広告を文書のページと関連付ける。本発明により的確に提供される広告は、たとえば、実質的に無料で広告が組み込まれる文書に関連付けられ、或いは、組み込まれた広告でもよい。



(2)

【特許請求の範囲】

【請求項1】 ユーザの入力したキーワードに基づいてハイパーテキスト文書を検索し、その結果を順位づけして表示する装置において、参照元付き文書表展開手段を備え、前記参照元付き文書表展開手段にあらかじめ各文書の参照元を表形式に展開しておくことで、参照元文書の情報を加味した順位づけ表示を可能とするハイパーテキスト検索装置。

【請求項2】 ハイパーテキスト文書管理手段によって管理されるハイパーテキストによる参照構造を持った文 10書群を解析し、各文書でとに当該文書を参照する参照元文書の識別子および参照元文書内のアンカー文を表形式にして参照元付き文書表として保持する参照元付き文書表展開手段と、前記参照元付き文書表内の各文書の情報に基づき索引を作成して保持する検索用索引展開手段と、ユーザからの検索要求を受け付けるキーワード入力手段と、前記検索用索引を調べ当該キーワードを含む文書群を検索する検索手段と、前記検索手段によって検索された文書群に該当する前記参照元付き文書表内の文書情報に基づき当該文書群に含まれる各文書の重要度合を 20算出し順位づけて表示データとする順位づけ生成手段と、当該順位づけ表示データをユーザに表示する検索結果表示手段から構成されるハイパーテキスト検索装置。

【請求項3】 前記順位づけ生成手段が、順位づけされる各文書の見出しに添えて、当該文書の参照元文書の見出しを前記参照元付き文書表内から取り出して選択表示データとすることを特徴とする請求項2記載のハイバーテキスト検索装置。

【請求項4】 前記順位づけ生成手段が、文書群の順位づけに際し前記参照元付き文書表を調べ、当該文書群内 30 のある文書と他の文書が近い参照関係にある場合、当該両文書を同一の順位にまとめた選択表示データを生成することを特徴とする請求項2記載のハイバーテキスト検索装置。

【請求項5】 前記順位づけ生成手段が、文書群の順位づけに際し前記参照元付き文書表を調べ、当該文書群内の各文書の重要度合に、当該文書の参照元のアンカー文の重要度合も加えて当該文書の重要度合として順位づけすることを特徴とする請求項2記載のハイパーテキスト検索装置。

【請求項6】 前記順位づけ生成手段が、順位づけされる各文書の見出しに添えて、前記参照元付き文書表内から当該文書の先頭部分およびキーワード周辺部分を取り出して選択表示データとすることを特徴とする請求項2記載のハイバーテキスト検索装置。

【請求項7】 前記順位づけ生成手段が、文書群の順位づけに際し前記参照元付き文書表を調べ、当該文書群内の各文書について当該文書に当該文書の参照元文書の内容を合わせて重要度合を算出し順位づけすることを特徴とする請求項2記載のハイパーテキスト検索装置。

【請求項8】 前記順位づけ生成手段が、文書群の順位づけに際し前記参照元付き文書表を調べ、当該文書群内の各文書について当該文書内の見出し部分の重要度合を他の部分よりも大きく評価して当該文書の重要度合として順位づけするととを特徴とする請求項2記載のハイパーテキスト検索装置。

【請求項9】 前記順位づけ生成手段が、文書群の順位づけに際し前記参照元付き文書表を調べ、当該文書群内の各文書の重要度合に、当該文書内で他の文書に対するアンカー文となっている部分の重要度合を減じて当該文書の重要度合として順位づけすることを特徴とする請求項2記載のハイバーテキスト検索装置。

【請求項10】 前記参照元付き文書表展開手段が前記 参照元付き文書表を作成するにあたり、一文書内に参照 ラベルが設定されている場合、当該文書を参照ラベルの 位置で分割して複数の文書として再定義することを特徴とする請求項2記載のハイバーテキスト検索装置。

【請求項11】 ハイパーテキスト文書管理手段によっ て管理されるハイパーテキストによる参照構造を持った 文書群を解析し、各文書ごとに当該文書を参照する参照 元文書の識別子および参照元文書内のアンカー文を表形 式にして参照元付き文書表として保持する参照元付き文 書表展開手段と、前記参照元付き文書表内の各文書の情 報に基づき索引を作成して保持する検索用索引展開手段 と、ユーザからの検索要求を受け付けるキーワード入力 手段と、前記検索用索引を調べ当該キーワードを含む文 書群を検索する検索手段と、前記検索手段によって検索 された文書群に該当する前記参照元付き文書表内の文書 情報に基づき当該文書群に含まれる各文書の重要度合を 算出し順位づけて表示データとする順位づけ生成手段 と、当該順位づけ表示データをユーザに表示する検索結 果表示手段から構成され、前記順位づけ生成手段が、順 位づけされた文書群の中に高い重要度合で現れる単語群 を同定し、当該単語群を前記キーワード入力手段に戻す ようにして、当該単語群から新たなキーワードを選択入 力することで再検索が起動されるようにしたことを特徴 とする請求項2記載のハイパーテキスト検索装置。

【請求項12】 前記順位づけ生成手段が、順位づけされた文書群の中の高い重要度合で現れる単語群を同定する際、キーワードの近傍に出現する単語の重要度合を他の部分よりも大きく評価して当該単語の重要度合とすることを特徴とする請求項11記載のハイバーテキスト検索装置。

【発明の詳細な説明】

[0001]

【発明の属する技術分野】本発明はハイパーテキスト検 索装置に関し、特に大量のハイバーテキスト文書の中か ら、ユーザの要求に合致しそうな文書を検索して提示す るハイパーテキスト検索装置に関する。

50 [0002]

【従来の技術】従来の、ユーザの要求に合致しそうな文 書を検索して提示する装置は、図11に示すような文書 検索装置10であった。検索用索引展開手段6は、文書管 理手段11に格納されている大量の文書をあらかじめ解析 し、内蔵の辞書に登録されている単語群内の各単語につ いて、当該単語がどの文書に何回現れているかを調べ、 当該単語の全文書における偏り(idf値)を求め、当 該単語が出現する各文書について規格化出現頻度(t f 値)を掛け合わせる t f×i d f 評価値を算出して検索 用索引を展開する。

【0003】キーワード入力手段2は、ユーザ9から複 数のキーワード入力を受け付け、検索手段3に供する。 検索手段3は、当該入力キーワードと合致する単語群を 前記検索用索引から探し、当該入力キーワードが出現す る文書の識別子と、当該文書における当該キーワードの t f×i d f評価値との対を複数個得て、順位づけ生成 手段4に供する。

【0004】順位づけ生成手段4は、tf×idf評価 値の高い順に当該文書群を整列させ、検索結果表示手段 5に供する。検索結果表示手段5は、前記整列順に文書 の識別タイトルを列挙したうえでユーザの選択指定を受 け付け、選択指定される都度、選択指定された文書のみ を文書管理手段11から呼び出して表示する。

【0005】以上のような構成の文書検索装置10は、ユ ーザが要求に沿ったキーワードを入力することで、大量 の文書からユーザの要求に近い文書をより近い順に提示 させることができる。

【0006】なお、tf×idf評価値の具体的な算出 方法は、いくつか知られているが、たとえば、当該単語 を含む文書数の全文書に対する割合を対数値にしたもの を1から引いた値をidf値とし、当該文書中の当該単 語の出現回数を当該文書の大きさで割った値をtf値と して両者を掛け合わせることで算出される。

【OOO7】tf×idf評価値の詳細および、それを 応用した文書検索装置については、「Gerard Salton: Introduction to Modern Information Retrieval, McGr aw-Hill, 1983」 に記載されている。

[0008]

【発明が解決しようとする課題】従来構成の文書検索装 置であっても、ハイパーテキスト文書の検索は不可能で 40 はない。しかしながら、ハイパーテキスト文書は各文書 が必ずしも独立でなく、参照関係にある文書を含めて初 めて一貫した内容となっていることも多い。そうしたハ イバーテキスト文書の検索を従来の文書検索装置で行な うと、ユーザにとっての文書の重要度合が誤って評価さ れて適切な順位づけができない問題や、順位づけされた 結果におけるユーザの選択指定が円滑に行なえない問題 が発生する。

【〇〇〇9】特に、ハイパーテキストマークアップ言語 (HTML) で書かれたインターネットのワールドワイドウ 50 と、前記検索手段によって検索された文書群に該当する

ェブ (www) 上の文書群では、一連の内容が参照関係に ある複数の文書にわたっている傾向が顕著で、問題が大

【0010】本発明は、こうした問題を解決し、ワール ドワイドウェブ (www) 上の文書群などの、これまで問 題が大きかったハイパーテキスト文書群に対しても検索 の利便性が高い文書検索装置を提供することを目的とす る。

[0011]

(3)

【課題を解決するための手段】本発明のハイパーテキス 10 ト検索装置は、あらかじめハイパーテキスト文書群を解 析して、各文書ごとに当該文書を参照する参照元文書の 識別子および参照元文書内のアンカー文を表形式にした 参照元付き文書表を展開しておくことで、参照元の情報 を加味した順位づけ表示を可能とするようにしたもので ある。

【0012】本発明によれば、検索および順位づけに際 して、各文書がどの文書によって参照されているかがハ イパーテキスト文書管理手段にいちいち問い合わせると となく判明するため、参照関係にある文書を統合した重 要度合の評価処理や、参照関係にある文書を統合した順 位づけ結果の表示処理を遅滞なく行なえる。その結果、 ユーザの要求に合致したより適切な順位づけや ユーザ の選択指定が容易な順位づけ表示が実現できる。 [0013]

【発明の実施の形態】本発明の請求項1記載の発明は ユーザの入力したキーワードに基づいてハイパーテキス ト文書を検索し、その結果を順位づけして表示する装置 において、参照元付き文書表展開手段を備え、前記参照 30 元付き文書表展開手段にあらかじめ各文書の参照元文書 を表形式に展開しておくことで、参照元の情報を加味し た順位づけ表示を可能とするハイバーテキスト検索装置 としたものであり、各文書がどの文書によって参照され ているかがハイパーテキスト文書管理手段にいちいち問 い合わせることなく、参照元付き文書表展開手段に展開 された参照元文書の表を引くだけで判明し、参照関係に ある文書群を統合した重要度合の評価処理や、参照関係 にある文書群を統合した順位づけ結果の表示データ生成 を高速に行なえるという作用を有する。

【0014】本発明の請求項2記載の発明は、ハイバー テキスト文書管理手段によって管理されるハイパーテキ ストによる参照構造を持った文書群を解析し、各文書と とに当該文書を参照する参照元文書の識別子および参照 元文書内のアンカー文を表形式にして参照元付き文書表 として保持する参照元付き文書表展開手段と、前記参照 元付き文書表内の各文書の情報に基づき索引を作成して 保持する検索用索引展開手段と、ユーザからの検索要求 を受け付けるキーワード入力手段と、前記検索用索引を 調べ当該キーワードを含む文書群を検索する検索手段

(4)

前記参照元付き文書表内の文書情報に基づき当該文書群 に含まれる各文書の重要度合を算出し順位づけて表示デ ータとする順位づけ生成手段と、当該順位づけ表示デー タをユーザに表示する検索結果表示手段から構成される ハイパーテキスト検索装置としたものであり、各文書が どの文書によって参照されているかがハイパーテキスト 文書管理手段にいちいち問い合わせることなく、参照元 付き文書表展開手段に展開された参照元文書の表を引く だけで判明し、参照関係にある文書群を統合した重要度 合の評価処理や、参照関係にある文書群を統合した順位 10 づけ結果の表示データ生成を高速に行なえるようになる とともに、ユーザが要求に沿ったキーワードを入力する ことで、大量のハイパーテキスト文書からユーザの要求 に近いハイパーテキスト文書群がより近い順に高速に提 示されるという作用を有する。

【0015】本発明の請求項3記載の発明は、前記順位 づけ生成手段が、順位づけされる各文書の見出しに添え て、当該文書の参照元文書の見出しを前記参照元付き文 書表内から取り出して選択表示データとすることを特徴 とする請求項2記載のハイパーテキスト検索装置とした ものであり、参照関係になっている文書群は通常深い関 連性がある内容を持っているので、検索された文書に関 連する文書のユーザによる選択呼び出しが容易となる作 用を有する。

【0016】本発明の請求項4記載の発明は、前記順位 づけ生成手段が、文書群の順位づけに際し前記参照元付 き文書表を調べ、当該文書群内のある文書と他の文書が 近い参照関係にある場合、当該両文書を同一の順位にま とめた選択表示データを生成することを特徴とする請求 項2記載のハイパーテキスト検索装置としたものであ り、参照関係になっている文書群は通常深い関連性があ る内容を持っているので、検索された文書群とそれらの 関連文書群の一覧性を高め、同時にユーザの同一文書の 重複呼び出しを防ぐという作用を有する。

【0017】本発明の請求項5記載の発明は、前記順位 づけ生成手段が、文書群の順位づけに際し前記参照元付 き文書表を調べ、当該文書群内の各文書の重要度合に、 当該文書の参照元のアンカー文の重要度合も加えて当該 文書の重要度合として順位づけすることを特徴とする請 求項2記載のハイパーテキスト検索装置としたものであ り、アンカー文は通常参照先の文書の特徴を良く記述し ているので、検索された文書群の順位づけの信頼性を高 めるという作用を有する。

【0018】本発明の請求項6記載の発明は、前記順位 づけ生成手段が、順位づけされる各文書の見出しに添え て、前記参照元付き文書表内から当該文書の先頭部分お よびキーワード周辺部分を取り出して選択表示データと することを特徴とする請求項2記載のハイパーテキスト 検索装置としたものであり、ユーザが順位づけされた各 文書の概要を当該文書を選択呼び出しすることなしに理 50 に含まれる各文書の重要度合を算出し順位づけて表示デ

解でき、ユーザによる検索結果の取捨選択が容易となる という作用を有する。

【0019】本発明の請求項7記載の発明は、前記順位 づけ生成手段が、文書群の順位づけに際し前記参照元付 き文書表を調べ、当該文書群内の各文書について当該文 書に当該文書の参照元文書の内容を合わせて重要度合を 算出し順位づけするととを特徴とする請求項2記載のハ イパーテキスト検索装置としたものであり、ユーザの検 索要求にかかわる内容が参照関係にある複数文書に分離 している場合に、当該複数文書が分離して扱われて順位 づけの低下を招くことを防ぐという作用を有する。

【0020】本発明の請求項8記載の発明は、前記順位 づけ生成手段が、文書群の順位づけに際し前記参照元付 き文書表を調べ、当該文書群内の各文書について当該文 書内の見出し部分の重要度合を他の部分よりも大きく評 価して当該文書の重要度合として順位づけすることを特 徴とする請求項2記載のハイバーテキスト検索装置とし たものであり、見出し部分が文書の特徴を通常良く記述 しているので、検索された文書群の順位づけの信頼性を 髙めるという作用を有する。

【0021】本発明の請求項9記載の発明は、前記順位 づけ生成手段が、文書群の順位づけに際し前記参照元付 き文書表を調べ、当該文書群内の各文書の重要度合に、 当該文書内で他の文書に対するアンカー文となっている 部分の重要度合を減じて当該文書の重要度合として順位 づけすることを特徴とする請求項2記載のハイパーテキ スト検索装置としたものであり、参照先ばかりが列挙さ れておりユーザが要求する内容が乏しい文書の順位づけ を低く抑えるという作用を有する。

【0022】本発明の請求項10記載の発明は、前記参 30 照元付き文書表展開手段が前記参照元付き文書表を作成 するにあたり、一文書内に参照ラベルが設定されている 場合、当該文書を参照ラベルの位置で分割して複数の文 書として再定義することを特徴とする請求項2記載のハ イパーテキスト検索装置としたものであり、内部に複数 の意味まとまりを持つ長大な文書を各意味まとまりごと に検索し順位づけを行なえるという作用を有する。

【0023】本発明の請求項11記載の発明は、ハイバ ーテキスト文書管理手段によって管理されるハイパーテ 40 キストによる参照構造を持った文書群を解析し、各文書 どとに当該文書を参照する参照元文書の識別子および参 照元文書内のアンカー文を表形式にして参照元付き文書 表として保持する参照元付き文書表展開手段と、前記参 照元付き文書表内の各文書の情報に基づき索引を作成し て保持する検索用索引展開手段と、ユーザからの検索要 求を受け付けるキーワード入力手段と、前記検索用索引 を調べ当該キーワードを含む文書群を検索する検索手段 と、前記検索手段によって検索された文書群に該当する 前記参照元付き文書表内の文書情報に基づき当該文書群

(5)

ータとする順位づけ生成手段と、当該順位づけ表示データをユーザに表示する検索結果表示手段から構成され、前記順位づけ生成手段が、順位づけされた文書群の中に高い重要度合で現れる単語群を同定し、当該単語群を前記キーワード入力手段に戻すようにして、当該単語群から新たなキーワードを選択入力することで再検索が起動されるようにしたことを特徴とするハイパーテキスト検索装置としたものであり、検索結果を見ながらの絞り込みを狙った再検索や興味の湧いたキーワードによる新たな検索を、キーボードを使うことなしにポインティング 10 デバイスのみを用いたキーワード入力が可能であるという作用を有する。

【0024】本発明の請求項12記載の発明は、前記順位づけ生成手段が、順位づけされた文書群の中に高い重要度合で現れる単語群を同定する際、当該入力キーワードの当該文書における出現位置の近傍に現れる単語の重要度合を他の部分に出現する単語よりも大きく評価して当該単語の重要度合とすることを特徴とする請求項11記載のハイパーテキスト検索装置としたものであり、キーワードの近傍に出現する単語が通常そのキーワードと関連性が高いので、関連キーワードとしての利用価値が高い単語群が選定列挙される可能性を高めるという作用を有する。

【0025】以下、本発明の実施の形態について、例として、インターネットのワールドワイドウェブ(www)上に分散管理されている文書群から、特定の文書群をキーワード検索するハイパーテキスト検索装置について図1から図10を用いて説明する。

【0026】一般にハイパーテキスト文書は図2のように、複数の参照関係にある文書80~86によって構成されている。文書80~84にはところどころにアンカー文800~806があり参照先の文書の識別子あるいは文書格納位置が埋め込まれている。ユーザはいわゆるブラウザと呼ばれる文書閲覧選択手段を用い、例えば文書81を閲覧しながらアンカー文801の位置をいわゆるボインティングデバイスで選択することにより、参照先の文書83を呼び出し閲覧するといった方法でハイパーテキスト文書を利用する。文書81に対して文書83を参照先文書、逆に文書83に対して文書81を参照元文書と呼ぶ。

【0027】ワールドワイドウェブWWW上のハイバーテキスト文書群はハイパーテキストマークアップ言語(HTML)で書かれており、ひとつの文書はページと呼ばれて、単に文字列だけでなく、画像やプログラムをも含んでいる。HTMLでは、nouka.htmlというファイルに格納された文書81内の「りんご生産農家」という文字列に対して、ringo.htmlというファイルに格納された文書83への参照を埋め込んでアンカー文とするには、文書81内の該当文字列部分を「りんご生産農家 のように表記すればよい。その場合、参照先文書83には何も記さなくて良いので、文書83が作成され

たあとに、文書83が格納されている計算機からはネットワーク上で遠く離れた位置にある計算機内に文書82が作成され、文書82が文書83の参照元文書のひとつになっているという場合も起こりうる。

【0028】(第1の実施の形態)図1は、本発明のハ イパーテキスト検索装置 1を含む全体構成を示すブロッ ク図である。図1のハイパーテキスト文書管理手段8 は、ワールドワイドウェブwwwの場合ではネットワーク 上に広域に分散された多数の計算機によって分散管理さ れている。そのため、本発明のハイパーテキスト検索装 置1の参照元付き文書表展開手段7は、関連文書収集機 能(一般的にウェブロボットと呼ばれる)を装備してお り、幾つかの文書格納位置アドレス(ユニバーサルリソ ースロケータ、URL)を与えると当該URLが指し示す文書 から順にアンカー文に埋め込まれているURLを手がかり に参照先の文書を次々と収集する。そして収集された文 書に、収集順などに基づく重複しない文書識別子を付 す。なお、画像やプログラムは除いて文字列のみを収集 すると収集時間を節約でき、また、すでに収集した文書 のURLをリストにしておき、当該リストにあるURLが指し 示す文書は重ねて収集しないこととすると、図2に示す 文書84から文書83を参照するアンカー文806のような戻 りの参照関係があるときに重複して文書を収集してしま うととが避けられる。

【0029】次に図1の参照元付き文書表展開手段7 は、収集した文書群から図3に示すような参照元付き文 書表を、次の手順で作成する。収集した文書数に相当す る個数の文書情報記入欄70~72を作成し、当該各文書情 報記入欄に文書識別子の番号と文書格納位置のURLアド 30 レスを記入する。収集した各文書について、当該HTML文 書内の文字列を調べ、「<title>うちのりんご</title >」などと記されている部分から「うちのりんご」など のタイトルを抽出して記入したうえ、BODY部に存在する 文字列から「<」と「>」で囲まれた部分であるHTMLタ グの文字列を除いてテキスト本文として記入する。収集 した各文書について、当該文書内の文字列を調べアンカ 一文があったならば、当該アンカー文の参照先文書に該 当する文書情報記入欄の参照元リストに、当該アンカー 文をもつ文書の識別子と当該アンカー文とを対にして記 40 入する。

【0030】とのように参照元付き文書表展開手段7は、ハイバーテキスト文書管理手段8に格納されている大量の文書をあらかじめ収集してそれを整理し、各文書ととに当該文書の識別子、文書格納位置、タイトル、テキスト本文などに加えて、当該文書を参照する参照元文書の識別子および参照元文書内のアンカー文を表形式にして参照元付き文書表として展開し、保持するものである。

家」 のように表記すればよい。その場合、参照先 【0031】なお、以上では参照元付き文書表の作成を 文書83には何も記さなくて良いので、文書83が作成され 50 全ての文書収集が終った後に行なうとして説明したが、 前記文書収集時にアンカー文を調べるのに合わせて、参 照元付き文書表の作成を順次行なう方法をとると効率化 が図れる。

【0032】また本発明の実施の形態におけるtf値 は、単語の文書内出現回数そのものを使用し、文書の大 きさで規格化しないとしている。これに対し文書の大き さで規格化する t f 値を使う場合は、各文書情報記入欄 に文書の大きさも算出して展開しておく必要がある。

【0033】図1の検索用索引展開手段6は、数万語の 辞書を有しており、図4に示すような検索用索引を、次 10 の手順で作成する。辞書に登録されている単語数に相当 する個数の単語情報記入欄60~62を作成し、当該単語情 報記入欄それぞれに単語を記入する。前記参照元付き文 書表内の全ての文書情報記入欄の内容を単語切りし、各 単語に相当する文字列が存在する文書の識別子と出現位 置(出現場所)を、該当する単語の単語情報記入欄に記

【0034】図4は、図3の文書83の内容を単語切りし た後の、「りんご」という単語の単語情報記入欄を示し ている。(TITLE,4)はタイトルの4文字目に「りん C」があること、(BODY,1,33,43) はテキスト本文の1 文字目、33文字目、43文字目に「りんご」があること、 (000082,5)は、文書82のアンカー文の5文字目に「り んじ」があることを意味する。また各単語の出現文書数 (あるいは i d f 値) や各文書内出現回数 (あるいは t f値)などのあらかじめ算出しておける数値は、検索用 索引展開時に算出して当該単語情報記入欄に記入してお くと、検索時の処理時間が節約できる。さらに各単語情 報記入欄に登録されている文書については、当該文書の 文書情報記入欄の単語リストに図3の最下段にあるよう に当該単語を並べておくと、文書情報から容易に単語情 報を引くことができる。

[0035] とのように検索用索引展開手段6は、前記 参照元付き文書表に格納されているテキスト本文、タイ トル、アンカー文などの文字列をあらかじめ解析し、内 蔵の辞書に登録されている単語群内の各単語について、 当該単語がどの文書内の文字列のどの位置に現れている かを記した検索用索引を展開し、保持するものである。 【0036】図1のキーワード入力手段2は、テキスト ボックス機能と当該テキストボックスの内容を返す検索 40 め込んでアンカー文としておく。 開始用ボタンを備えて「検索のページ」などと題したHT ML文書を用いて実現される。ユーザは、自分のクライア ント計算機で起動したモザイクあるいはネットスケープ などのwwwブラウザで当該HTML文書を呼び出し、テキス トボックスに単一のキーワードを入力するかあるいは複 数のキーワードをスペースで区切って列挙して入力し、 続けて検索開始用ボタンを押すことにより、キーワード の入力がなされる。

【0037】このようにキーワード入力手段2は、ユー

に供給するものである。なお、キーワードの候補となる 複数の単語をボタンとして列挙し選択入力可能としてお くと、キーボードから任意文字列を入力するのが不得手 のユーザにも、当該キーワード候補の単語範囲で、キー ボードを使わずともポインティングデバイスを用いたキ ーワード入力が可能である。

【0038】図1の検索手段3は、前記キーワード入力 手段2によって入力されたキーワード群と同一単語の単 語情報記入欄を前記検索用索引から取り出す。

【0039】とのように検索手段3は、当該入力キーワ ードと合致する単語群を前記検索用索引から探し、当該 入力キーワードが出現する文書の識別子と、当該キーワ ードが出現する当該文書の文字列内の位置情報との組を 複数個得て、順位づけ生成手段4に供するものである。 【0040】図1の順位づけ生成手段4は、当該キーワ ードが出現する各文書について t f×i d f 評価値を算 出して当該文書の重要度合とし、当該重要度合が高い順 に当該文書群を順位づける。入力されたキーワードが複 数ある場合は、各々のキーワードに関して出現する文書 のtf×idf評価値を算出して重要度合とし、全文書 にわたり重要度合を比較して順位づける。

【0041】なお、複数種のキーワードを含む文書につ いて各キーワードのtf×idf評価値を足し合わせて 当該文書の重要度合として順位づける。また、複数種の キーワードを含む文書について t f×i d f 評価値を足 し合わせる際に、キーワードを2種含む時はさらに2倍 するとか、キーワードを3種含む時はさらに3倍すると かの加点(重み付け)を行なうとキーワードの共起によ る順位づけが上がり、ユーザの要求に沿うことが多くな る。さらに2種のキーワードが文書内で20文字より近い 距離にある場合はさらに2倍するなどの、キーワードの 文書内近接共起による加点(重み付け)を行なうと、キ ーワード共起による順位づけがさらに強く効き、ユーザ の要求に沿うことが多くなる。

【0042】順位づけ生成手段4は、さらに、順位づけ された順で各文書の見出しを列挙したHTML文書を表示用 に作成し、検索結果表示手段5に供する。見出しとは、 各文書のタイトルあるいはその参照元文書のアンカー文 の文字列であり、個々の見出しには当該文書のURLを埋

【0043】とのように順位づけ生成手段4は、前記参 照元付き文書表から参照関係にある文書群を導き、当該 文書群を統合した重要度合の評価処理および順位づけ並 びに順位づけ結果の表示データ生成を行ない、検索結果 表示手段5に供するものである。

【0044】図1の検索結果表示手段5は、ユーザのク ライアント計算機で起動したモザイクあるいはネットス ケープなどのwwwブラウザで実現され、表示用に作成さ れたHTML文書を受け付けクライアント計算機に画面表示 ザ9から複数のキーワード入力を受け付け、検索手段3 50 すると同時に、ポインティングデバイスで見出しを選択

指定することで、見出しに対応する実体の文書が当該見 出しに埋め込まれているURLに基づいて呼び出し可能と

【0045】とのように検索結果表示手段5は、前記表 示データを画面表示したうえでユーザの選択指定を受け 付け、選択指定される都度、選択指定された文書のみを ハイパーテキスト文書管理手段8から呼び出して表示す るものである。

【0046】以上のように本発明の第1の実施の形態に よれば、各文書がどの文書によって参照されているかが 10 ハイパーテキスト文書管理手段8にいちいち問い合わせ ることなく、参照元付き文書表展開手段7に展開された 参照元文書の表を引くだけで判明し、参照関係にある文 書群を統合した重要度合の評価処理や、参照関係にある 文書群を統合した順位づけ結果の表示データ生成を高速 に行なえるようになるとともに、ユーザが要求に沿った キーワードを入力することで、大量のハイパーテキスト 文書からユーザの要求に近いハイパーテキスト文書群が より近い順に高速に提示されるという有利な効果を奏す

【0047】(第2の実施の形態)本発明の第2の実施 の形態は、前記第1の実施の形態と基本的な構成は同じ である。しかし、ハイパーテキスト検索装置1の順位づ け生成手段4の機能を若干異にしている。たとえば図2 に示される参照関係にある文書群から文書83、文書85、 文書86が検索され、各々4位、18位、19位に順位づけら れた場合を考える。

[0048]順位づけ生成手段4は、順位づけ結果生成 時に、検索された各文書(図6の星印を付した文書)の 参照元文書を参照元付き文書表を用いてたどっていき、 たとえば1つ上と2つ上の参照元文書の見出しを添えた 順位づけ結果表示が、図6のように生成できる。文書83 の参照元関係にある文書は文書81と文書82の複数がある が、文書格納位置の部分文字列同士をマッチさせること で、より文書格納位置が近い文書を判定して参照元文書 をひとつに絞っている。とのように参照元文書の見出し を添えておけば、図6の結果表示から、19位にある文書 86のようにアンカー文をひとつも持たない文書(図2参 照)を選択し呼び出してしまっても、再び図6の結果表 示に戻れば、深い関連性がある内容をもつであろう参照 40 元の文書83や文書81が容易に選択呼び出し可能となる。 【0049】以上のように本発明の第2の実施の形態に よれば、順位づけ生成手段4において、前記参照元付き 文書表から検索された文書の参照元関係にある文書群の 見出しを取り出し、順位づけ結果の表示データ生成の際 に、検索された文書の見出しだけでなく当該参照元文書 の見出しも合わせて表示データとし、検索結果表示手段 5に供することにより、参照関係になっている文書群は 通常深い関連性がある内容を持っているので、検索され

易となる有利な効果を奏する。

【0050】(第3の実施の形態)本発明の第3の実施 の形態は、前記第1の実施の形態と基本的な構成は同じ である。しかし、ハイパーテキスト検索装置1の順位づ け生成手段4の機能を若干異にしている。また前記第2 の実施の形態で行なった図6のような表示では問題が多 少残ることに対して考慮したものである。たとえば図2 および図6において、4位にある文書83を呼び出して読 み、さらに文書83にあるアンカー文から文書85を呼び出 して読んだとする。その後に、5位、6位と読み進む と、すでに読み終っている文書85が18位に出現するので 読み終っていると気づかずに再度呼び出してしまった り、19位で文書86を呼び出して読んでも、文書83を読ん でから時間がたっているので文脈がよく理解できなかっ たりする可能性がある。

【0051】そこで、本発明の第3の実施の形態では、 図1の順位づけ生成手段4において、前記参照元付き文 書表から検索された文書の参照元関係にある文書群の識 別子を取り出し他の検索された文書の識別子と同一性を 判定することで、検索された文書同士が近い参照関係に あることを認定し、順位づけ結果の表示データ生成の際 に、近い参照関係が認定された当該文書群の見出しを同 一の順位の表示データにして検索結果表示手段5に供す るようにしている。その結果、図6の4位の文書群と18 位の文書群と19位の文書群のように共通の文書を含む文 書群は、それらの文書識別子から判定して、図8のよう に4位に一体化して表示するようにすると、検索された 文書83、文書85、文書86が参照関係の文脈を保持しなが ら続いて読むことができるので、ユーザによる検索結果 30 の一覧性を高めると同時に、ユーザの同一文書の重複呼 び出しを防ぐことができる。

【0052】なお、以上では一体化表示時に、一体化さ れる文書のうち最も重要度合の大きな文書の重要度合 を、一体化された文書群の重要度合として順位づけると して説明したが、一体化された文書群の重要度合を、一 体化された文書群の重要度合の和として順位づけしても よい。

【0053】(第4の実施の形態)本発明の第4の実施 の形態は、前記第1の実施の形態と基本的な構成は同じ である。しかし、ハイパーテキスト検索装置1の順位づ け生成手段4の機能を若干異にしている。 すなわち、本 発明の第4の実施の形態では、たとえば入力キーワード 「りんご」に関する検索では、図4の「りんご」の単語 情報記入欄が指し示している「りんご」の出現位置が文 書81および文書82のアンカー文であっても、当該出現回 数2回を文書83のテキストに出現した「りんご」の出現 回数4回に加えて6回としidf×tf評価値を計算し て順位づけると、アンカー文では通常参照先の文書の特 徴を記述したキーワード(この場合は「りんご」)が何 た文書に関連する文書のユーザによる選択呼び出しが容 50 度も現れるので、検索された文書83の順位づけが高くな

14

り、「りんご」に関する検索の信頼性を高められる。 【0054】つまり、第4の実施の形態では、図1の順位づけ生成手段4において、検索手段3から供された入力キーワードの出現位置が前記参照元付き文書表内の参照元文書のアンカー文であっても、当該出現回数を参照先の文書内のテキストに出現した当該入力キーワードの出現回数に加えて当該参照先文書の重要度合を計算して順位づけたうえで表示データを生成し、検索結果表示手段5に供するようにする。これにより、アンカー文は通常参照先の文書の特徴を良く記述しているので、検索された文書群の順位づけの信頼性を高める有利な効果を奏する。

【0055】(第5の実施の形態)本発明の第5の実施の形態は、前記第1の実施の形態と基本的な構成は同じである。しかし、ハイパーテキスト検索装置1の順位では生成手段4の機能を若干異にしている。すなわち、本発明の第5の実施の形態では、順位では結果生成時に、単語情報記入欄から得られる入力キーワードのテキスト本文内の出現位置にもとづき、当該入力キーワードを含む文あるいはフレーズを抜き出して当該文書のテキスト本文の先頭文あるいは先頭フレーズに続けて表記した要約文を形成し、図7に示すように、各文書の見出しに添えて表示データとすると、ユーザが順位でけされた文書をいちいち選択呼び出しすることなしに、順位でけ結果表示画面に表記された各文書の要約文を読むだけで当該文書の概要を理解でき、ユーザによる検索結果の取捨選択が容易となる。

【0056】つまり、第5の実施の形態では、図1の順位づけ生成手段4において表示データを生成するに際し、検索手段3から供された入力キーワードの前記参照 30元付き文書表内の出現位置にもとづき、当該入力キーワードを含む文あるいはフレーズを抜きたして当該文書のテキスト本文の先頭文あるいは先頭フレーズに添えて表示データとし、検索結果表示手段5に供するようにする。これにより、ユーザが順位づけされた各文書の概要を当該文書を選択呼び出しすることなしに理解でき、ユーザによる検索結果の取捨選択が容易となる有利な効果を奏する。

【0057】なお、以上ではひとつの文書にキーワードが何度現れても当該入力キーワードを含む文あるいはフレーズを全て抜きだして要約文を生成するとして説明したが、5種までの複数キーワード検索は1種当たり3箇所までしか抜き出さないとか、10種までの複数キーワード検索は1種当たり2箇所までしか抜き出さないとか、10種を越える複数キーワード検索は1種当たり1箇所までしか抜き出さないとかの制限を設けると、要約文が長大になって順位づけの一覧性が低下するのを防ぐことができる。

[0058] (第6の実施の形態)本発明の第6の実施 表内の見出し部分の文字列にあるとき、当該出現回数をの形態は、前記第1の実施の形態と基本的な構成は同じ 50 数倍したうえで文書内のテキスト全体に出現した当該入

である。しかし、ハイパーテキスト検索装置1の順位づ け生成手段4の機能を若干異にしている。すなわち、ワ ールドワイドウェブwwwの文書(ページ)では、各文書 が細切れになっており、一文書内に十分な文脈が記述さ れていないことが多い。たとえば文書83の「りんご」が 「青森りんご」であるにもかかわらず、「青森」の記述 は文書83にはなく参照元の文書81にあることがある。 【0059】本発明の第6の実施の形態では、たとえば 入力キーワード「りんご」と「青森」に関する検索にお 10 いて文書83の重要度合を算出するにあたり、文書83と参 照元文書81を合わせた文書の重要度合と、文書83と参照 元文書82を合わせた文書の重要度合とを比較し、「青 森」という単語が入っていて重要度合が大きい文書81と 文書83を合わせた文書の重要度合を文書83の重要度合と すると、文書83の順位づけの低下を防ぐことができる。 【0060】つまり、第6の実施の形態では、図1の順 位づけ生成手段4において、検索手段3によって検索さ れた文書群内の各文書について前記参照元付き文書表か ら当該文書の参照元関係にある文書群を特定し、当該検 索された文書の内容と当該参照元関係にある文書の内容 を複合した内容で重要度合を算出し、当該参照元関係に ある文書群のうち当該検索された文書と複合して最も評 価が高くなる文書を選び、当該検索された文書の重要度 合に代えて当該複合文書の重要度合を用いて順位づけし たうえで表示データとし、検索結果表示手段5に供する ようにする。これにより、ユーザの検索要求にかかわる 内容が参照関係にある複数文書に分離している場合に、

【0061】(第7の実施の形態)本発明の第7の実施の形態は、前記第1の実施の形態と基本的な構成は同じである。しかし、ハイパーテキスト検索装置1の順位でけ生成手段4の機能を若干異にしている。すなわち、本発明の第7の実施の形態では、入力キーワード「りんで」の出現位置が文書のタイトルあるいは参照元文書のアンカー文などの見出し部分の文字列にあるとき、文書83内のテキスト全体に出現した当該入力キーワードの出現回数6回のうちアンカー文とタイトルに出現した回数3回を2倍にし、出現回数を9回として当該文書の重要度合を計算するなどとすると、見出し部分が文書の特徴を通常良く記述しているので、文書83の順位づけの信頼性をさらに高められる。なお、見出し部分にはアンカー文とタイトルに加え、HTML文では「<hl>」と「</hl>
「で囲まれる文字列に代表される小見出しを含めても良い

当該複数文書が分離して扱われて順位づけの低下を招く

ことを防ぐという有利な効果を奏する。

【0062】つまり、本発明の第7の実施の形態では、図1の順位づけ生成手段4において、検索手段3から供された入力キーワードの出現位置が前記参照元付き文書表内の見出し部分の文字列にあるとき、当該出現回数を数倍したうえで文書内のテキスト全体に出現した当該入

50

力キーワードの出現回数に加えて当該文書の重要度合を 算出し、当該文書を順位づけたうえで表示データを生成 し、検索結果表示手段5に供するようにする。これによ り、見出し部分が文書の特徴を通常良く記述しているの で、検索された文書群の順位づけの信頼性を高める有利 な効果を奏する。

15

【0063】(第8の実施の形態)本発明の第8の実施 の形態は、前記第1の実施の形態と基本的な構成は同じ である。しかし、ハイパーテキスト検索装置1の順位づ け生成手段4の機能を若干異にしている。すなわち、ワ 10 ールドワイドウェブwwwの文書(ページ)には、他の文 書を参照するアンカー文が多数列挙されているいわゆる リンクページが存在する。リンクページは検索表示され ても検索意図に合った有益な情報がないことが多い。

【0064】本発明の第8の実施の形態では、たとえば 文書82のテキスト本文が「りんご」に関するアンカー文 10個のみのリンクページであり「りんご」の出現回数が テキスト本文で10回であった場合でも、「りんど」に関 する検索において文書83での「りんご」の出現回数を調 べるにあたり、参照元文書のアンカー文に「りんご」が 20 見い出される都度、文書82のテキスト本文での「りん ど」の出現回数を1回分減じることを繰り返すことによ り、文書82のテキスト本文での「りんご」の出現回数を 0回にまで低下させられ、リンクページの順位づけを低 く抑えることができる。

【0065】つまり、本発明の第8の実施の形態では、 図1の順位づけ生成手段4において、検索手段3から供 された入力キーワードの出現位置が前記参照元付き文書 表内の参照元文書のアンカー文であるとき、当該出現回 数を当該参照元の文書内のテキスト本文に出現した当該 入力キーワードの出現回数から減じて当該参照元文書の 重要度合を計算して順位づけたうえで表示データを生成 し、検索結果表示手段5に供するようにする。こうする ことにより、参照先ばかりが列挙されておりユーザが要 求する内容が乏しい文書の順位づけを低く抑えることが できる。

・【0066】 (第9の実施の形態) 本発明の第9の実施 の形態は、前記第1の実施の形態と基本的な構成は同じ である。しかし、ハイパーテキスト検索装置1の参照元 付き文書表展開手段7の機能を若干異にしている。すな 40 わち、本発明の第9の実施の形態では、参照元付き文書 表として展開するに際し、図5の文書87のように「<an ame="ABC">」などのラベルがあり、かつ当該文書内から あるいは他の文書から「」などと参照 されている場合には、当該文書を参照ラベルの位置の直 前で分割し、ラベル以降を各々別の文書(文書88、文書 89) として扱って参照元付き文書表を作っておくと、内 部に複数の意味まとまりを持つ長大な文書を各意味まと まりごとに検索し順位づけを行なえる。なお、上に示さ れるような参照ラベルに加えて、「<h1>」と「</h1>」

で囲まれる文字列に代表される小見出しの位置の直前で もって分割し、各々別の文書として扱っても良い。

【0067】つまり、本発明の第9の実施の形態では、 図1の参照元付き文書表展開手段7において、ハイパー テキスト文書管理手段8に格納されている大量の文書を あらかじめ収集整理し参照元付き文書表として展開する に際し、各文書に当該文書内からあるいは他の文書から の参照ラベルが付されているか判定し、参照ラベルが付 されている場合には当該文書を参照ラベルの位置で分割 し、各々別の文書として扱って参照元付き文書表を作っ ておき、前記検索用索引展開およびキーワード入力によ る検索と順位づけを行なうようにする。こうすることに より、内部に複数の意味まとまりを持つ長大な文書を各 意味まとまりごとに検索し順位づけを行なえるという有 利な効果を奏する。

【0068】(第10の実施の形態)図9は、図1の本 発明のハイパーテキスト検索装置1と若干異なるハイパ ーテキスト検索装置91を含む全体構成を示すブロック図 である。図9の構成と図1の構成とで異なる構成は、図 9の順位づけ生成手段94の順位づけ結果がキーワード入 力手段92に戻されている点である。その他の構成は、図 1と同じなので構成の説明は省略することにする。

【0069】本発明の第10の実施の形態では、順位づ け結果生成時に、順位づけされた文書群内のたとえば上 位10文書について、各文書の文書情報記入欄の単語リス トを調べ、そとに存在する各単語について当該10文書内 のtf×idf評価値を算出して各単語の重要度合と し、重要度合の高い順にたとえば10単語を選んでキーワ ード入力手段92に戻して、当該10単語を選択ボタンとし たHTML文書を生成する。そして図10に示すように、検 索結果表示手段5は、wwwブラウザのネットスケープの フレーム機能を用いて表示画面を2分割し、画面上段に 順位づけ結果表示窓50を画面下段に関連キーワード入力 窓20を設け当該HTML文書を表示することにより、たとえ ば「りんご」の検索結果を見ながら、より検索意図に近 い「特産品」で絞り込みを狙った再検索を起動したり、

「ねぶた」などの興味の湧いた他のキーワードによる新 たな検索を起動するのが、キーボードを使うことなしに ポインティングデバイスのみを用いた入力で可能であ る。

【0070】なお、図10における「特産品」「ねぶ た」などの単語ボタンは当該単語をテキストボックスに 追加するJAVAスクリプトを動作させ、「クリア」のボタ ンはテキストボックスの単語群をクリアするJAVAスクリ プトを動作させ、「初期状態」のボタンはテキストボッ クスの単語群を前回の検索に使用したキーワード群 (と の場合は「りんご」)に戻すJAVAスクリプトを動作さ せ、「再検索」ボタンでテキストボックスの単語群をキ ーワードとして検索が再起動されるように実装される。 【0071】なお、以上の説明における順位づけされた

文書群内の上位10文書とは、検索された文書(図8また は図10の星印の文書)を想定したが、検索された文書 の参照元文書(図8または図10の順位づけ表示に現れ る文書群)までに広げて上位10位までに現れる全文書す ると、参照関係にある文書から広く関連キーワードが収 集できる。

【0072】つまり、第10の実施の形態では、図9の 順位づけ生成手段94において、前記参照元付き文書表か ら順位づけされた文書群に存在する単語を調べ、各単語 についての重要度合を前記検索用索引の情報から算出 し、重要度合の高い順に数単語をキーワード入力手段92 に戻して選択ボタンにした表示を、検索結果表示手段5 の表示と同一画面内に配置するようにする。こうすると とにより、検索結果を見ながらの絞り込みを狙った再検 索や興味の湧いたキーワードによる新たな検索を、キー ボードを使うことなしにポインティングデバイスのみを 用いたキーワード入力が可能であるという有利な効果を 奏する。

【0073】(第11の実施の形態)本発明の第11の 実施の形態は、前記第10の実施の形態と基本的な構成 20 は同じである。しかし、ハイパーテキスト検索装置91の 順位づけ生成手段94の機能を若干異にしている。すなわ ち、本発明の第11の実施の形態では、関連キーワード の重要度合を算出するにあたり、各単語について前記検 索用索引の情報から文書中の出現位置を調べ、当該入力 キーワードの前後20文字などの近傍位置に出現した単語 は、その出現回数を2倍にしたうえでtf×idf評価 値を求めて重要度合とし、その高い順に10単語を選定し て関連キーワード入力窓20に順に表示すると、図10に 示すように、「りんご」という入力キーワードの近傍に 30 出現しやすく「りんと」に高い関連性を持つ「信州」

「農家」などの単語が、入力キーワードの近傍には出現 しにくいが関連文書にはよく出現する「長野」「オリン ピック」などの単語よりも先に表示される。これによ り、絞り込みを狙った再検索に使う単語が最初の方にあ り、新たな検索の興味をひく単語が後の方にある傾向を もった単語群の列挙表示が可能となり、当該単語群の関 連キーワードとしての利用価値が高まる。

【0074】つまり、第11の実施の形態では、図9の 順位づけ生成手段94において、前記参照元付き文書表か 40 20 関連キーワード入力窓 ら順位づけされた文書群に存在する単語群を調べ、各単 語についての重要度合を前記検索用索引の情報から算出 する際、当該キーワードの近傍に出現した単語の重要度 合を数倍したうえで、重要度合の高い順に数単語を選定 し列挙するようにする。とうすることにより、キーワー

ドの近傍に出現する単語が通常そのキーワードと関連性 が高いので、関連キーワードとしての利用価値が高い単 語群が選定列挙される可能性を高めることができる。 [0075]

【発明の効果】以上のように本発明によれば、参照関係 にある文書を統合した重要度合の評価処理や、参照関係 にある文書を統合した順位づけ結果の表示処理を効率良 く行なえ、ユーザの要求に合致したより適切な順位づけ や、ユーザの選択指定が容易な順位づけ表示が実現でき 10 るという有利な効果が得られる。

【図面の簡単な説明】

- 【図1】本発明の全体構成を示すブロック図、
- 【図2】ハイパーテキスト文書の構造を示す図、
- 【図3】収集した文書群から作成された参照元付き文書 表を示す図、

【図4】参照元付き文書表内の全ての文書情報記入欄の 内容を単語切りし、各単語毎に作成される検索用索引の 一例を示す図、

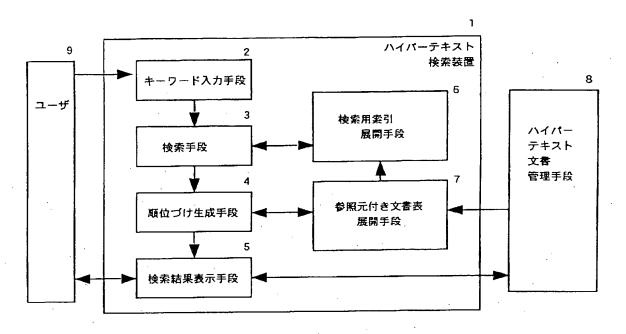
- 【図5】ラベル付き文書の分割の様子を示す図。
- 【図6】順位づけ結果を表示する第1の図、
 - 【図7】順位づけ結果を表示する第2の図、
 - 【図8】順位づけ結果を表示する第3の図、
 - 【図9】関連キーワードによる再検索機能を実現する構 成を示すブロック図、

【図10】関連キーワード付き順位づけ結果を表示する

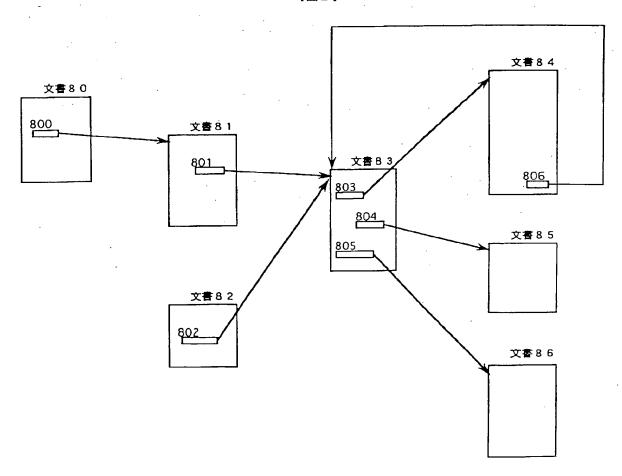
【図11】従来例の全体構成を示すブロック図である。 【符号の説明】

- 1、91 ハイパーテキスト検索装置
- 2、92 キーワード入力手段
 - 3 検索手段
 - 4、94 順位づけ生成手段
 - 5 検索結果表示手段
 - 6 検索用索引展開手段
 - 7 参照元付き文書表展開手段
 - 8 ハイパーテキスト文書管理手段
 - 9 ユーザ
 - 10 文書管理手段
 - 11 文書検索装置
- - 50 順位づけ結果表示窓
 - 60~62 単語情報記入欄
 - 70~72 文書情報記入欄
 - 80~89 文書
 - 800~806 アンカー文

[図1]



【図2】



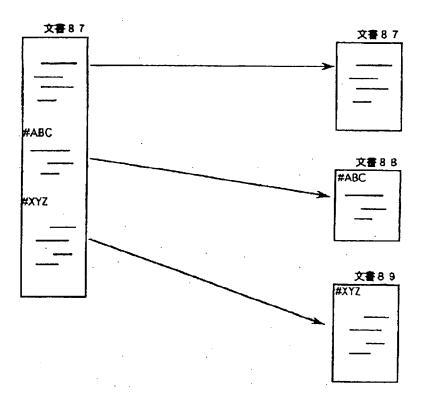
[図3]

		70 / 1	
文書識別子	0 0 0 0 8 3		
文書格納位置	www.mei.co.jp/home/tokusan/aomori/ringo.html		
タイトル・	うちのりんご		
テキスト本文	りんごがたくさんとれま おいしそうでしょう。 農園の場所案内。 りんごの作り方。 宅配りんご、注文の仕り	·	
参照元リスト	参照元文書識別子	アンカー文	
	000081	りんご生産農家	
	000082	おいしいりんごもあります。	
単語リスト	りんご おいし 展園	場所 案内 作り方 宅配 注文 仕方 生産 農家	

【図4】

		60 / /
単語	りんご	
出現文書リスト	出现文書識別子	文書内出現場所
	000083	(TITLE, 4) (BODY, 1, 33, 43) (000081, 1) (000082, 5)
	· ·	
	:	

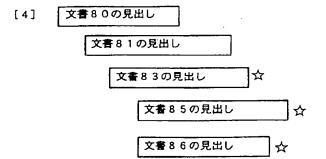
【図5】



【図8】

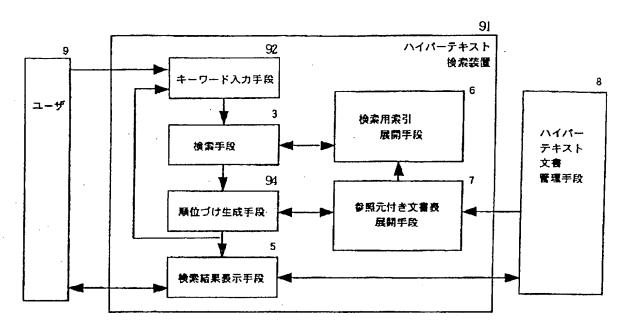
検索結果

検索キー [りんご] 86件中30件表示

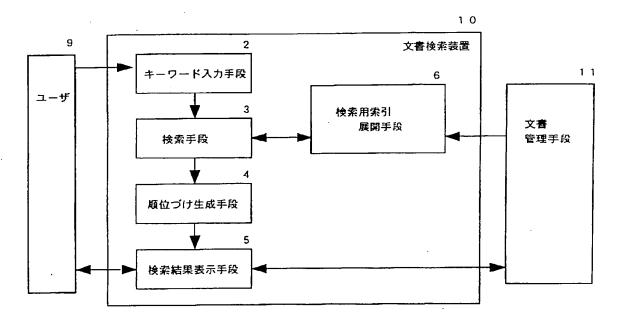


[図7] 【図6】 検索結果 検索結果 検索キー [りんご] 検索キー [りんご] 86件中30件表示 86件中30件表示 [4] 文書80の見出し [4] 文書80の見出し 文書81の見出し 文書81の見出し 文書83の見出し ☆ 文書83の見出し 文書83の要約文 。 [18] 文書81の見出し 文書83の見出し [18] 文書81の見出し 文書85の見出し ☆ 文書83の見出し [19] 文書81の見出し 文書85の見出し 쇼 文書83の見出し 文書85の要約文 文書86の見出し ☆ [19] 文書81の見出し 文書83の見出し 文書8.6の見出し ☆ 文書86の要約文

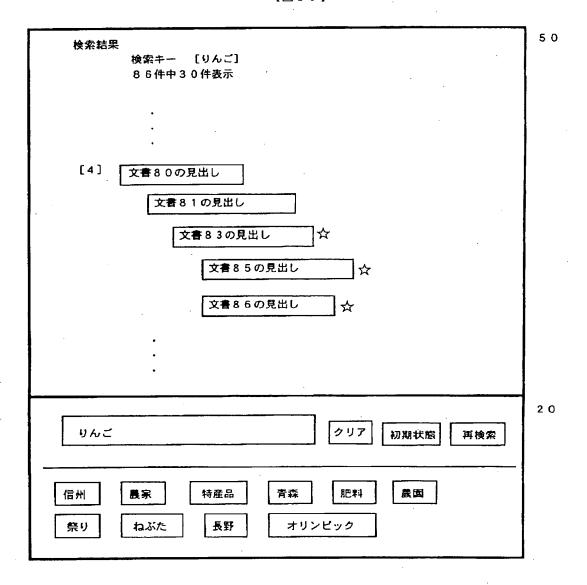
【図9】



【図11】



【図10】



フロントページの続き (51)Int.Cl.⁶ 識別記号 庁内整理番号 FΙ 技術表示箇所 G06F 17/21 G 0 6 F 15/20 590E 15/40 370A 15/403 370Z (72)発明者 野口 喜洋 (72)発明者 安川 秀樹 大阪府門真市大字門真1006番地 松下電器 大阪府門真市大字門真1006番地 松下電器 産業株式会社内 産業株式会社内